*Economics and Philosophy*, **20** (2004) 19–33 Copyright © Cambridge University Press
DOI: 10.1017/S0266267104001270

# EQUAL VALUE OF LIFE AND THE PARETO PRINCIPLE

ANDREAS HASMAN* AND LARS PETER ØSTERDAL†
**University of Oxford*
†*University of Copenhagen*

A principle claiming equal entitlement to continued life has been strongly defended in the literature as a fundamental social value. We refer to this principle as 'equal value of life'. In this paper we argue that there is a general incompatibility between the equal value of life principle and the weak Pareto principle and provide proof of this under mild structural assumptions. Moreover we demonstrate that a weaker, age-dependent version of the equal value of life principle is also incompatible with the weak Pareto principle. However, both principles can be satisfied if transitivity of social preference is relaxed to quasi-transitivity.

## 1. INTRODUCTION

Recent years have seen vigorous debate about ethical principles for resource allocation in health care. In this paper we outline a general framework that enables us to investigate possible implications and limitations of a principle of 'equal value of life'. This principle claims equal entitlement to continued life and has received notable support in both the philosophical and the health economics literature.

The equal value of life principle originates from the position that interpersonal comparison of utility is either meaningless or unethical and that it is therefore wrong to discriminate on the basis of health state. Accepting this premise we therefore consider social welfare orderings of life profiles that do not necessarily make use of interpersonal utility

The authors are grateful to Luc Bovens, Kristian Schultz Hansen, Søren Holm, Franz Huber, Wiebke Kuklys, Gabriella Pigozzi, and two anonymous referees for detailed and very helpful comments on earlier versions of this paper. Any errors and shortcomings are the responsibility of the authors alone.

comparisons. We nevertheless find that any social welfare ordering of life profiles satisfying the equal value of life principle violates the weak Pareto principle. The same is the case for a weaker, age-dependent version of the principle. In other words, *all* individuals lose out from adherence to the (age-dependent) equal value of life principle. In fact, even individuals at less desirable health states, who are otherwise meant to be protected by the principle, will have good reason not to support it. A conflict with collective rationality must indeed be of concern to those defending the equal value of life principle on grounds of distributive justice.

After establishing this we turn to possible alternative ways of validating the equal value of life principle. Many people would probably find that the principle, at least in an age-dependent version, has some intuitive appeal. Referring to intuition alone is not satisfactory, however, and in order to defend the notion of equal value of life we should be able to identify well-described situations of social choice where the principle is consistent with rational social decision-making. We find that under quasi-transitivity, where strict social preference is transitive, the equal value of life principle is in fact consistent with the Pareto principle. If both principles are satisfied, however, there is no ordering of social states and social preferences cannot be represented by a social welfare function.

The remainder of the paper is organised as follows. In sections 2 and 3 we introduce and discuss an 'equal value of life' principle and consider Pareto conditions for social choice. We provide the necessary definitions in section 4, and formulate incompatibilities between equal value of life principles and the weak Pareto principle in section 5. In section 6 we consider a possible resolution to these incompatibilities and conclude in section 7.

## 2. EQUAL VALUE OF LIFE

The value of human lives could be the most important consideration when decisions are made on claims for health care resources. Under resource constraints one objective could be to maximise the over-all value of life derived from health care provision. When comparing the relative value of lives, it seems possible to discriminate on a number of counts, e.g. quality of life or expected utility for particular groups. Some authors fiercely reject differentiated weighing of individual lives, however, and assert that continued life has equal value for everyone. In this paper we consider a fixed population and discuss one possible formulation of the 'equal value of life' principle, namely that which holds the claim that a gain in life years for one individual is equally as good as the same gain in life years for another individual. We restrict the principle to situations where the individuals prefer the potential gain in lifetime to immediate death. We also consider an age-dependent restricted version of this principle.

Harris is among the strongest adherents to a concept of equal value of life. He argues that the value of life depends exclusively on a capacity in the individual to value her/his own life and those of others:

> If we allow that the value of life for each individual consists simply in those reasons, *whatever they are*, that each person has for finding their own life valuable and for wanting to go on living, then we do not need to know what the reasons are. All we need to know is that particular individuals have their own reasons, or rather, simply, that they value their own lives. (Harris 1985: 16).

This argument clearly leads to the conclusion that lives are in fact equally valuable, even if some lives are not lived at perfect health, as long as they are valued by those living those lives. Harris argues further that

> while it follows from the fact (if it is a fact) that I and everyone else would prefer to have, say one year of healthy life rather than three years of severe discomfort, that we value healthy existence more than uncomfortable existence for ourselves, it does not follow that where the choice is between three years of discomfort for *me* or immediate death on the one hand, and one year of health for *you* on the other, that I am somehow committed to the judgement that you ought to be saved rather than me.[1] (Harris 1987: 118)

From this it appears that the equal value of life principle has two main elements. Firstly it asserts that it is wrong to base resource allocation decisions on a comparison of people's utility from health services and, secondly, it expresses an ethical concern for a fundamental entitlement to continued life, held equally and to the same extent by all persons.[2] The principle of 'equal value of life' has also gained some support among health economists. Nord (2001: 580), for example, applies similar reasoning when arguing that

> providers of health utilities come across as making statements about the lack of well being . . . of people with health problems without having asked them. Worse than that, users of health state utilities are effectively saying that the value of saving a disabled person's life is lower than the value of saving a healthy person's life. Both these statements are insulting.

The rejection of interpersonal comparisons of health state utility has been questioned by other authors however. Based on a Harsanyian veil-of-ignorance style argument Singer *et al.* (1995) conclude that valuing lives unequally, and maximising aggregated value derived from health care,

[1] Harris compares gains in life years that are not necessarily equally long. Our definition of 'equal value of life' is restricted in the sense that it only applies to situations where two different persons can gain the *same* number of life years.

[2] Since we consider a *fixed* identifiable population of individuals, each being alive at least until they are born, we avoid some of the difficulties with welfare comparisons of populations that vary in size and identity of individuals (see e.g. Arrhenius 2000, and references therein).

is in fact *not* wrong as doing so would be rational in situations where members of society are unaware of their own future states of health (see also McKie *et al.* 1998). Savulescu (1998) questions the reliability of Harris' clearly subjective account for the value of life. According to Savulescu it cannot be up to the individual to decide the value of her or his life if this value is to serve as a general principle for the allocation of scarce resources. Williams (1997) also disagrees with Harris when he infers that the value of lives legitimately can be distinguished based on a combination of life expectancy and quality of life so that the lives of those who will benefit the most are valued more highly when resources are allocated.[3]

At the centre of the debate about the value of life appears to be a conflict between two very diverse ideas: on the one hand an equal value of life principle which stresses the value of equal individual entitlement to continued life and rejects interpersonal utility comparisons and, on the other hand, methods based on aggregation of standardised health utilities, such as maximisation of the total number of quality-adjusted life years (QALYs) in the population, where preference intensity has meaning and utility trade-offs are made explicitly between persons.

We will argue that although an equal value of life principle may be intuitively appealing, it comes into conflict with the requirements of rational social decision-making; but this point has nothing to do with interpersonal utility comparisons.

## 3. PARETO CONDITIONS AND SOCIAL CHOICE

The notion of Pareto conditions on social choice is a cornerstone in welfare economic theory. This theory entails that people have individual preferences over some set of conceivable alternatives (here: individual health factors). Some alternatives are preferred over others and therefore have higher individual value. Social welfare is defined over the distribution of individual alternatives in society and the Pareto principle is a social value judgement pertaining to these distributions. The principle occurs in a weak and a strong form. The weak Pareto principle states that: *If a change is beneficial for everyone in society then it is a change for the better*, and in the strong version: *If a change is beneficial for at least one member of society and worse for nobody else then it is a change for the better* (Shaw 1999). To some commentators the Pareto principle is self-evident and obviously rational, although this view has been challenged by some authors (see e.g. Culyer and Wagstaff 1993; Cohen 1995; Shaw 1999).

[3] See the debates in *Bioethics* (Harris 1996, 1999; Savulescu 1998, 1999), *British Medical Journal* (Culyer 1997; Harris 1997; Williams 1997), *Health Economics* (Johannesson 2001; Nord 2001; Williams 2001) and *Journal of Medical Ethics* (Harris 1987, 1995; Savulescu 1995; Singer *et al.* 1995).

It is clear that neither version of the Pareto principle involves interpersonal comparisons of preference intensities, and in this respect the underlying premise of the Pareto principles resembles that behind the equal value of life principle. Thus, what we need to know to establish adherence to the Pareto principles is simply whether a social intervention leaves anyone better off or no one worse off and not how persons compare to each other (Shaw 1999: 362).

Here we focus on the weak Pareto principle for two principal reasons. First, the strong Pareto principle may not seem as convincing as the weaker version. The fact that *some* individuals are left better off by a social intervention might not trump concerns for equality if society's primary objective is to maximise the well-being of the worst off individual. One could thus argue that acknowledging the social value of making *all* people better off has more direct intuitive appeal as a necessary condition for rational social decision-making. Second, a violation of the weak Pareto principle tells us more than a violation of the strong version of the principle: if we find the strong version violated, but the weak version possibly satisfied, it reveals only a minor weakness in social welfare assessment (assuming that Paretian reasoning is relevant at all). A violation of the weak version, however, indicates a more fundamental problem.

In a recent paper, Kaplow and Shavell (2001) showed that if individual preferences and the social welfare ordering are represented by continuous real-valued functions, then the weak Pareto principle implies that the social welfare function depends only on individual utility levels. Although we do not assume here that individual and social preferences necessarily have a real-valued representation, our work is related in the sense that we also find that the weak Pareto principle gives rise to quite strong restrictions on social welfare assessment.

Culyer and Wagstaff (1993) also investigated the incompatibility of various notions of individual entitlements in the particular context of health care resource allocation. They considered a model where individual health is represented by a real-valued health index to which individual entitlement is directly related. This is a utility measure, which abstracts from underlying determinants of health. However, there are potentially many ways of aggregating underlying health factors into an individual health index, and each method will ultimately lead to different social choices. As we discussed earlier, the principle of 'equal value of life' is not related to interpersonal utility comparisons, but pertains to a more fundamental equal entitlement to years of life. In order to analyse this we do not take individual utility levels as the primitive, but consider more generally *life profiles*, which are paths specifying individual health states from birth to death. Social welfare is accordingly evaluated over distributions of life profiles in the society.

## 4. NOTATION AND ASSUMPTIONS

In the following, let $A$ denote the set of conceivable health states for an individual. We might think of these health states as being defined in terms of a variety of different aspects of health such that they give a complete description of the well-being of the person. $A$ is thus an abstract set with no particular mathematical structure. The health state 'dead' is contained in $A$ and is denoted $a_0$. Let $t_i$ be a non-negative finite number that denotes the age of person $i$ when he or she dies. If $t_i = 0$ we say that individual $i$ has a *zero-life* with the interpretation that he or she dies at the moment he or she is born.[4] If an individual lives for $t_i$ years, a health state different from $a_0$ is experienced in the time up to $t_i$ after which the individual remains at $a_0$. Accordingly, a *life profile* is represented by a map $l_i : \mathbb{R}_+ \to A$, where $l_i(s)$ is the health state of individual $i$ at age $s$, $l_i(s) \neq a_0$ for $0 \leq s < t_i$, and $l_i(s) = a_0$ for $s \geq t_i$. Let $L$ denote the collection of all conceivable life profiles, and let $\succsim_i$ be a complete and transitive binary relation on $L$ representing individual $i$'s preferences for own life profiles.[5,6] Strict preference $\succ_i$ and indifference $\sim_i$ are defined from $\succsim_i$ in the usual way.

For any individual $i$ we assume that there exist health states $\overline{a}_i, \underline{a}_i \in A$, different from $a_0$, where individual $i$ always prefers health state $\overline{a}_i$ to $\underline{a}_i$ in the sense that replacing health state $\underline{a}_i$ with $\overline{a}_i$ for an interval of lifetime with positive length makes individual $i$ strictly better off. Moreover, we assume that $\overline{a}_i$ and $\underline{a}_i$ are better than death in the sense that individual $i$ always prefers extending a life at health states $\overline{a}_i$ or $\underline{a}_i$ to $a_0$ (death).

We also assume a restricted form of *continuity* of individual preferences between health states $\overline{a}_i$, $\underline{a}_i$ and $a_0$. Let $l_i$ and $l_i'$ be life profiles and let $l_i[\varepsilon]$ and $l_i'[\varepsilon]$ be life profiles that agree with $l_i$ and $l_i'$ respectively except that within an interval of lifetime of length $\varepsilon$ health states $\overline{a}_i$, $\underline{a}_i$ or $a_0$ can replace each other. Then if $l_i \succ_i l_i'$ we assume that there is $\varepsilon > 0$ sufficiently small such that $l_i[\varepsilon] \succ_i l_i'[\varepsilon]$ for any $l_i[\varepsilon]$ and $l_i'[\varepsilon]$.

[4] Of course, other interpretations are possible. A special case is where $t_i$ is interpreted as the number of years the individual is alive after conception. In this case, we may want to restrict attention to positive lifetimes since the zero-life would not have a clear interpretation in this context.

[5] $l_i \succsim_i l_i'$ is read as 'life profile $l_i$ is at least as good as life profile $l_i'$ for individual $i$'. A binary relation $\succsim_i$ is *complete* if for all pairs $l_i, l_i' \in L$ that $l_i \succsim_i l_i'$ or $l_i' \succsim_i l_i$, and *transitive* if for all triples $l_i, l_i', l_i'' \in L$ that $l_i \succsim_i l_i'$ and $l_i' \succsim_i l_i''$ implies $l_i \succsim_i l_i''$.

[6] A special case is the QALY (quality-adjusted life years) model where individual health-related utility is assumed to be represented by a function of the form

$$q(l_i) = \int_0^{t_i} u(l_i(s))\, ds,$$

where $u$ measures instantaneous utility and $q$ measures overall utility of a life profile. Some form of discounting is occasionally added to the model. (Mild technical assumptions are necessary to obtain a well-defined integral.)

https://doi.org/10.1017/S0266267104001270 Published online by Cambridge University Press

We may, for example, think of $\overline{a}_i$ as a state with good health and $\underline{a}_i$ as a state with moderate health problems. Since the assumptions above are made only for individual preferences regarding own health they seem largely unobjectionable. Note also that we do not assume that individual preference can be represented by a utility function on $L$.

For a society with $n$ individuals, indexed $N = \{1, \ldots, n\}$, a *social preference relation* is a binary relation on the collection of all distributions of life profiles. Thus, we write $(l_1(\cdot), \ldots, l_n(\cdot)) \succsim (l'_1(\cdot), \ldots, l'_n(\cdot))$ if the distribution $(l_1(\cdot), \ldots, l_n(\cdot))$ is at least as socially desirable as $(l'_1(\cdot), \ldots, l'_n(\cdot))$. A *social welfare ordering* is a complete and transitive social preference relation.

## 5. INCOMPATIBILITIES

A social preference relation satisfies *equal value of life* if for any distribution of life profiles $s$ extra life years to individual $i$ is of equal social value to $s$ extra life years to another individual $j$, as long as the respective gains are better than death for each individual's own point of view. Formally, for a distribution $(l_1, \ldots, l_n)$ with life years $(t_1, \ldots, t_n)$ then for all $i, j \in N, i \neq j$ and $s > 0$

$$\left(l_i^{t_i+s}, l_{k\in N\setminus\{i\}}\right) \sim \left(l_j^{t_j+s}, l_{k\in N\setminus\{j\}}\right),$$

where $l_k^{t_k+s}$ is a life profile that agrees with $l_k$ up to the first $t_k$ life years with $l_k^{t_k+s} \succ_k l_k$, $k = i, j$.

A social preference relation satisfies the *weak Pareto principle* if $(l_1, \ldots, l_n) \succ (l'_1, \ldots, l'_n)$ when $l_i \succ_i l'_i$ for all $i$, and it satisfies the *strong Pareto principle* if $(l_1, \ldots, l_n) \succ (l'_1, \ldots, l'_n)$ when $l_i \succsim_i l'_i$ for all $i$ and $l_i \succ_i l'_i$ for at least one $i$.

We can now formulate the first main observation: The equal value of life principle and the weak Pareto principle are mutually exclusive. Once the framework has been formulated this result, stated in Proposition 1 below, is not difficult to establish. Yet, it illustrates a central point which informs later discussion.

**Proposition 1.** *There exists no social welfare ordering satisfying equal value of life and the weak Pareto principle. Moreover, if all individuals prefer any life with positive lifetime to the zero-life then a social welfare ordering satisfying equal value of life depends only on the total number of life years in the population.*

Proof. Consider a distribution $(l_{i\in N}^{t_i})$ where $l_i^{t_i}$ is a life profile for individual $i$ with $t_i$ life years which is preferred to the zero-life if $t_i > 0$. For individual 1, for any $s > 0$ let $l_1^{t_1+s}$ be a life with $t_1 + s$ life years that agrees with $l_1^{t_1}$ the first $t_1$ years, i.e. $l_1^{t_1+s}(r) = l_1^{t_1}(r)$ for $0 \le r \le t_1$, and with $l_1^{t_1+s} \succ_1 l_1^{t_1+s'}$ for any $s > s' \ge 0$. Note that such life profile can always be defined, for example with $l_1^{t_1+s}(r) = \overline{a}_1$ for all $t_1 \le r \le t_1 + s$.

https://doi.org/10.1017/S0266267104001270 Published online by Cambridge University Press

Let $T = t_1 + \ldots + t_n$, and for an individual $i$ let $l_i^0$ be a zero-life. We claim that $(l_{i\in N}^{t_i}) \sim (l_1^T, l_{k\in\{2,\ldots,n\}}^0)$. For this, let $2 \leq m \leq n$ and consider the distribution

$$\left(l_1^{t_1+t_2+\ldots+t_{m-1}}, l_{i\in\{2,\ldots,m\}}^0, l_{i\in\{m+1,\ldots,n\}}^{t_i}\right).$$

Then by equal value of life

(1) $$\left(l_1^{t_1+t_2+\ldots+t_m}, l_{i\in\{2,\ldots,m\}}^0, l_{i\in\{m+1,\ldots,n\}}^{t_i}\right) \sim \left(l_1^{t_1+t_2+\ldots+t_{m-1}}, l_{i\in\{2,\ldots,m-1\}}^0, l_{i\in\{m,\ldots,n\}}^{t_i}\right).$$

Since (1) holds for any $m$ by transitivity of $\sim$ we obtain

(2) $$\left(l_{i\in N}^{t_i}\right) \sim \left(l_1^T, l_{i\in\{2,\ldots,n\}}^0\right).$$

Let $l_1^T$ and $m_1^T$ be life profiles with $T$ life years for individual 1 which are preferred to the zero-life, and let $l_2^T$ be a life profile for individual 2 preferred to the zero-life. Then by equal value of life we have $(l_1^T, l_{i\in\{2,\ldots,n\}}^0) \sim (l_2^T, l_{i\in\{1,3,4,\ldots,n\}}^0)$ and $(m_1^T, l_{i\in\{2,\ldots,n\}}^0) \sim (l_2^T, l_{i\in\{1,3,4,\ldots,n\}}^0)$. Thus, by transitivity

(3) $$\left(l_1^T, l_{i\in\{2,\ldots,n\}}^0\right) \sim \left(m_1^T, l_{i\in\{2,\ldots,n\}}^0\right).$$

Combining (2) and (3) we find that if all individuals prefer any life with positive lifetime to the zero-life, then for any two distributions $(l_{i\in N}^{t_i})$ and $(m_{i\in N}^{t_i'})$ we have $(l_{i\in N}^{t_i}) \sim (m_{i\in N}^{t_i'})$ if $t_1 + \ldots + t_n = t_1' + \ldots + t_n'$, i.e. $\succsim$ depends only on the total number of life years in the population.

Now, to provoke a violation of the weak Pareto principle, let $(\bar{l}_{i\in N}^{t_i})$ be a distribution where each individual $i$ has $t_i$ life years at health state $\bar{a}_i$ and let $(\underline{l}_{i\in N}^{t_i'})$ be a distribution where each individual has $t_i'$ life years at health state $\underline{a}_i$ (see section 4). From (2) and (3) we have $(\bar{l}_{i\in N}^1) \sim (\underline{l}_{i\in N}^1)$. ■

Proposition 1 indicates that any method for allocating health care resources inevitably violates the equal value of life or the weak Pareto principle.[7] There may be reasonable arguments for rejecting interpersonal utility comparisons, such as those put forward by Harris and Nord. However, impossibility of interpersonal utility comparisons does not necessarily imply justification for equal value of life. As shown, the contrary seems to be the case: Even in a world where no meaning is attached to interpersonal utility comparisons, there is a strong argument against the equal value of life principle.

[7] Note that the continuity assumption for individual preference is not used in the proof.

Moreover, if the equal value of life principle holds, then two distributions of life profiles which are identical with respect to the distribution of life years and where each life profile is better than the zero-life are socially equally desirable, even if the first distribution throughout contains life profiles that are better than the life profiles in the second distribution. Most people would probably find this absurd. It is therefore natural to search for a weakening of the equal value of life principle. A weaker version states that equal value of life only applies between individuals at the same age. A social preference relation satisfies *age-dependent equal value of life* if for any distribution of life profiles where two individuals $i$ and $j$ live for $t$ years that $s$ extra life years to individual $i$ is of equal social value to $s$ extra life years to individual $j$, as long as the respective gains are better than death from each individual's own point of view. Formally, for a distribution $(l_1, \ldots, l_n)$ with life years $(t_1, \ldots, t_n)$ then for all $i, j \in N, i \neq j$ and $s > 0$

$$\left(l_i^{t+s}, l_{k \in N \setminus \{i\}}\right) \sim \left(l_j^{t+s}, l_{k \in N \setminus \{j\}}\right),$$

where $t = t_i = t_j$ and where $l_k^{t_k+s}$ is a life profile that agrees with $l_k$ up to the first $t_k$ life years with $l_k^{t_k+s} \succ_k l_k, k = i, j$.

The question is now whether an age-dependent equal value of life principle can be reconciled with Paretian welfare maximisation. Unfortunately, it cannot.

**Proposition 2.** *There exists no social welfare ordering satisfying age-dependent equal value of life and the weak Pareto principle.*

The proof is divided into four steps. In step 1, by repeated use of age-dependent equal value of life and transitivity we obtain indifference between two particular distributions which are identical besides certain improvements for all individuals apart from individual 1. Step 2 obtains equivalence of two other distributions which are identical besides a certain improvement for individual 1. In step 3 some implications of continuity (for individual 1) are derived, and finally in step 4 the pieces are put together to provoke a contradiction with the weak Pareto principle.

Let $(a[t], a'[t'], \ldots)_i$ be notation for a life profile with the first $t$ years in health state $a$ then $t'$ years in health state $a'$, etc., followed by death. When necessary we use a subscript $i$ to indicate that the life profile is for individual $i$. Let $\overline{a}_i$ and $\underline{a}_i$ be health states as defined in section 4. To ease readability, in step 1 and 2 we occasionally use overbrackets to point to the difference between two distributions.

Proof. Assume that the social welfare ordering $\succsim$ satisfies age-dependent equal value of life and the weak Pareto principle.

**Step 1.** For $0 \leq x \leq 1$, $\tau > 0$ and $2 \leq k \leq n$ consider the distribution

$$\{(\underline{a}_1[xt], \overline{a}_1[(1-x)t])_1, (\underline{a}_2[t], \underline{a}_2[\tau])_2, \ldots, (\underline{a}_{k-1}[t], \underline{a}_{k-1}[\tau])_{k-1}, (\underline{a}_k[t])_k, (\underline{a}_{k+1}[t], \overline{a}_{k+1}[\tau])_{k+1}, \ldots, (\underline{a}_n[t], \overline{a}_n[\tau])_n\}.$$

Note that individual 1 lives $xt$ years in $\underline{a}_1$ and then $(1-x)t$ years in $\overline{a}_1$, for $2 \leq h \leq k-1$ individual $h$ lives $t+\tau$ years in $\underline{a}_h$, individual $k$ lives $t$ years in $\underline{a}_k$, and for $h \geq k+1$ individual $h$ lives $t$ years in $\underline{a}_h$ followed by $\tau$ years in $\overline{a}_h$. Then by age-dependent equal value of life

$$\begin{aligned}&\{(\underline{a}_1[xt], \overline{a}_1[(1-x)t], \overbrace{\underline{a}_1[\tau]})_1, (\underline{a}_2[t], \underline{a}_2[\tau])_2, \ldots, (\underline{a}_{k-1}[t], \underline{a}_{k-1}[\tau])_{k-1},\\ &\quad (\underline{a}_k[t])_k, (\underline{a}_{k+1}[t], \overline{a}_{k+1}[\tau])_{k+1}, \ldots, (\underline{a}_n[t], \overline{a}_n[\tau])_n\}\\ &\sim \{(\underline{a}_1[xt], \overline{a}_1[(1-x)t])_1, (\underline{a}_2[t], \underline{a}_2[\tau])_2, \ldots, (\underline{a}_{k-1}[t], \underline{a}_{k-1}[\tau])_{k-1},\\ &\quad (\underline{a}_k[t], \overbrace{\underline{a}_k[\tau]})_k, (\underline{a}_{k+1}[t], \overline{a}_{k+1}[\tau])_{k+1}, \ldots, (\underline{a}_n[t], \overline{a}_n[\tau])_n\}\end{aligned}$$

and

$$\begin{aligned}&\{(\underline{a}_1[xt], \overline{a}_1[(1-x)t], \overbrace{\underline{a}_1[\tau]})_1, (\underline{a}_2[t], \underline{a}_2[\tau])_2, \ldots, (\underline{a}_{k-1}[t], \underline{a}_{k-1}[\tau])_{k-1},\\ &\quad (\underline{a}_k[t])_k, (\underline{a}_{k+1}[t], \overline{a}_{k+1}[\tau])_{k+1}, \ldots, (\underline{a}_n[t], \overline{a}_n[\tau])_n\}\\ &\sim \{(\underline{a}_1[xt], \overline{a}_1[(1-x)t])_1, (\underline{a}_2[t], \underline{a}_2[\tau])_2, \ldots, (\underline{a}_{k-1}[t], \underline{a}_{k-1}[\tau])_{k-1},\\ &\quad (\underline{a}_k[t], \overbrace{\overline{a}_k[\tau]})_k, (\underline{a}_{k+1}[t], \overline{a}_{k+1}[\tau])_{k+1}, \ldots, (\underline{a}_n[t], \overline{a}_n[\tau])_n\},\end{aligned}$$

hence by transitivity

$$\begin{aligned}&\{(\underline{a}_1[xt], \overline{a}_1[(1-x)t])_1, (\underline{a}_2[t], \underline{a}_2[\tau])_2, \ldots, (\underline{a}_{k-1}[t], \underline{a}_{k-1}[\tau])_{k-1},\\ &\quad (\underline{a}_k[t], \overbrace{\underline{a}_k[\tau]})_k, (\underline{a}_{k+1}[t], \overline{a}_{k+1}[\tau])_{k+1}, \ldots, (\underline{a}_n[t], \overline{a}_n[\tau])_n\}\\ &\sim \{(\underline{a}_1[xt], \overline{a}_1[(1-x)t])_1, (\underline{a}_2[t], \underline{a}_2[\tau])_2, \ldots, (\underline{a}_{k-1}[t], \underline{a}_{k-1}[\tau])_{k-1},\\ &\quad (\underline{a}_k[t], \overbrace{\overline{a}_k[\tau]})_k, (\underline{a}_{k+1}[t], \overline{a}_{k+1}[\tau])_{k+1}, \ldots, (\underline{a}_n[t], \overline{a}_n[\tau])_n\}.\end{aligned}$$

Since this holds for any $k$, $2 \leq k \leq n$, by transitivity of $\sim$ we have

$$\begin{aligned}(4)\quad &\{(\underline{a}_1[xt], \overline{a}_1[(1-x)t])_1, (\underline{a}_2[t], \underline{a}_2[\tau])_2, \ldots, (\underline{a}_n[t], \underline{a}_n[\tau])_n\}\\ &\sim \{(\underline{a}_1[xt], \overline{a}_1[(1-x)t])_1, (\underline{a}_2[t], \overline{a}_2[\tau])_2, \ldots, (\underline{a}_n[t], \overline{a}_n[\tau])_n\}.\end{aligned}$$

**Step 2.** Now, consider the distribution

$$\{(\underline{a}_1[t], \underline{a}_1[\tau])_1, (\underline{a}_2[t], \underline{a}_2[\tau/2], \overline{a}_2[\tau/2])_2, \ldots, (\underline{a}_n[t], \underline{a}_n[\tau/2], \overline{a}_n[\tau/2])_n\}.$$

https://doi.org/10.1017/S0266267104001270 Published online by Cambridge University Press

Then by age-dependent equal value of life

$$\begin{aligned}&\{(\underline{a}_1[t], \underline{a}_1[\tau], \overbrace{\underline{a}_1[\tau]})_1, (\underline{a}_2[t], \underline{a}_2[\tau/2], \overline{a}_2[\tau/2])_2, \ldots,\\ &\qquad (\underline{a}_n[t], \underline{a}_n[\tau/2], \overline{a}_n[\tau/2])_n\}\\ &\quad\sim \{(\underline{a}_1[t], \underline{a}_1[\tau])_1, (\underline{a}_2[t], \underline{a}_2[\tau/2], \overline{a}_2[\tau/2], \overbrace{\underline{a}_2[\tau]})_2, \ldots,\\ &\qquad (\underline{a}_n[t], \underline{a}_n[\tau/2], \overline{a}_n[\tau/2])_n\}\end{aligned}$$

and

$$\begin{aligned}&\{(\underline{a}_1[t], \underline{a}_1[\tau], \overbrace{\overline{a}_1[\tau]})_1, (\underline{a}_2[t], \underline{a}_2[\tau/2], \overline{a}_2[\tau/2])_2, \ldots,\\ &\qquad (\underline{a}_n[t], \underline{a}_n[\tau/2], \overline{a}_n[\tau/2])_n\}\\ &\quad\sim \{(\underline{a}_1[t], \underline{a}_1[\tau])_1, (\underline{a}_2[t], \underline{a}_2[\tau/2], \overline{a}_2[\tau/2], \overbrace{\underline{a}_2[\tau]})_2, \ldots,\\ &\qquad (\underline{a}_n[t], \underline{a}_n[\tau/2], \overline{a}_n[\tau/2])_n\},\end{aligned}$$

thus by transitivity

$$\begin{aligned}(5)\quad&\{(\underline{a}_1[t], \underline{a}_1[\tau], \overbrace{\underline{a}_1[\tau]})_1, (\underline{a}_2[t], \underline{a}_2[\tau/2], \overline{a}_2[\tau/2])_2, \ldots,\\ &\qquad (\underline{a}_n[t], \underline{a}_n[\tau/2], \overline{a}_n[\tau/2])_n\}\\ &\quad\sim \{(\underline{a}_1[t], \underline{a}_1[\tau], \overbrace{\overline{a}_1[\tau]})_1, (\underline{a}_2[t], \underline{a}_2[\tau/2], \overline{a}_2[\tau/2])_2, \ldots,\\ &\qquad (\underline{a}_n[t], \underline{a}_n[\tau/2], \overline{a}_n[\tau/2])_n\}.\end{aligned}$$

**Step 3.** Let $0 \leq x \leq 1$ and let $(\underline{a}_1[xt], \overline{a}_1[(1-x)t])$ be a life for individual 1 as in step 1.

If $x = 0$ then by continuity there is $2\tau^* > 0$ sufficiently small such that

$$(\overline{a}_1[t]) = (\underline{a}_1[xt], \overline{a}_1[(1-x)t]) \succ_1 (\underline{a}_1[t], \underline{a}_1[\tau^*], \overline{a}_1[\tau^*]).$$

Let $\tau^*$ be fixed. We claim that there is some $0 < x^* < 1$ such that

$$(6)\quad (\underline{a}_1[t], \underline{a}_1[\tau^*], \overline{a}_1[\tau^*]) \succ_1 (\underline{a}_1[x^*t], \overline{a}_1[(1-x^*)t]) \succ_1 (\underline{a}_1[t], \underline{a}_1[\tau^*], \underline{a}_1[\tau^*]).$$

For this, suppose that this is *not* the case and define

$$S \equiv \{x \mid (\underline{a}_1[xt], \overline{a}_1[(1-x)t]) \succsim_1 (\underline{a}_1[t], \underline{a}_1[\tau^*], \overline{a}_1[\tau^*])\}.$$

From the case $x = 0$ it is clear that $S$ is non-empty. Let $\overline{x} = \sup S$.

If $x = 1$, since $\underline{a}_1$ is preferred to death for individual 1, we have

$$\begin{aligned}&(\underline{a}_1[t], \underline{a}_1[\tau^*], \overline{a}_1[\tau^*]) \succ_1 (\underline{a}_1[t], \underline{a}_1[\tau^*], \underline{a}_1[\tau^*]) \succ_1 (\underline{a}_1[xt], \overline{a}_1[(1-x)t])\\ &\quad = (\underline{a}_1[t]).\end{aligned}$$

By continuity there is $\epsilon > 0$ such that

$$(\underline{a}_1[t], \underline{a}_1[\tau^*], \overline{a}_1[\tau^*]) \succ_1 (\underline{a}_1[xt], \overline{a}_1[(1-x)t]),$$

https://doi.org/10.1017/S0266267104001270 Published online by Cambridge University Press

for all $1-\epsilon \leq x \leq 1$. Hence $\overline{x} < 1$. From the case $x = 0$ in a similar way we can show that $\overline{x} > 0$. Thus we have $0 < \overline{x} < 1$.

Now, if

$$(\underline{a}_1[\overline{x}t], \overline{a}_1[(1-\overline{x})t]) \succsim_1 (\underline{a}_1[t], \underline{a}_1[\tau^*], \overline{a}_1[\tau^*])$$

then (because $l_1 \succsim_1 l_1'$ and $l_1' \succ_1 l_1''$ implies $l_1 \succ_1 l_1''$) we have

$$(\underline{a}_1[\overline{x}t], \overline{a}_1[(1-\overline{x})t]) \succ_1 (\underline{a}_1[t], \underline{a}_1[\tau^*], \underline{a}_1[\tau^*])$$

and

$$(\underline{a}_1[t], \underline{a}_1[\tau^*], \underline{a}_1[\tau^*]) \succsim_1 (\underline{a}_1[xt], \overline{a}_1[(1-x)t])$$

for all $x > \overline{x}$ contradicting continuity. On the other hand, if

$$(\underline{a}_1[t], \underline{a}_1[\tau^*], \underline{a}_1[\tau^*]) \succsim_1 (\underline{a}_1[\overline{x}t], \overline{a}_1[(1-\overline{x})t])$$

then (because $l_1 \succ_1 l_1'$ and $l_1' \succsim_1 l_1''$ implies $l_1 \succ_1 l_1''$) we have

$$(\underline{a}_1[t], \underline{a}_1[\tau^*], \overline{a}_1[\tau^*]) \succ_1 (\underline{a}_1[\overline{x}t], \overline{a}_1[(1-\overline{x})t])$$

and

$$(\underline{a}_1[xt], \overline{a}_1[(1-x)t]) \succsim_1 (\underline{a}_1[t], \underline{a}_1[\tau^*], \overline{a}_1[\tau^*])$$

for all $x < \overline{x}$ contradicting continuity and proving the claim.

**Step 4.** From (6) and the weak Pareto principle we now have

$$\begin{aligned}&\{(\underline{a}_1[t], \underline{a}_1[\tau^*], \overline{a}_1[\tau^*])_1, (\underline{a}_2[t], \underline{a}_2[\tau^*/2], \overline{a}_2[\tau^*/2])_2, \ldots,\\ &\quad (\underline{a}_n[t], \underline{a}_n[\tau^*/2], \overline{a}_n[\tau^*/2])_n\}\\ &\quad \succ \{(\underline{a}_1[x^*t], \overline{a}_1[(1-x^*)t])_1, (\underline{a}_2[t], \underline{a}_2[\tau^*])_2, \ldots, (\underline{a}_n[t], \underline{a}_n[\tau^*])_n\}.\end{aligned}$$

But then by (4) and (5) (and because $(l_1, \ldots, l_n) \succ (l_1', \ldots, l_n')$, $(l_1, \ldots, l_n) \sim (l_1'', \ldots, l_n'')$ and $(l_1', \ldots, l_n') \sim (l_1''', \ldots, l_n''')$ implies $(l_1'', \ldots, l_n'') \succ (l_1''', \ldots, l_n''')$) we have

$$\begin{aligned}&\{a_1[t], \underline{a}_1[\tau^*], \underline{a}_1[\tau^*])_1, (\underline{a}_2[t], \underline{a}_2[\tau^*/2], \overline{a}_2[\tau^*/2])_2, \ldots,\\ &\quad (\underline{a}_n[t], \underline{a}_n[\tau^*/2], \overline{a}_n[\tau^*/2])_n\}\\ &\quad \succ \{(\underline{a}_1[x^*t], \overline{a}_1[(1-x^*)t])_1, (\underline{a}_2[t], \overline{a}_2[\tau^*])_2, \ldots, (\underline{a}_n[t], \overline{a}_n[\tau^*])_n\}\end{aligned}$$

which contradicts the weak Pareto principle. ■

## 6. QUASI-TRANSITIVE SOCIAL PREFERENCE

Proving Propositions 1 and 2 we made use of transitivity of social indifference. Whereas transitivity of the strict social preference component

https://doi.org/10.1017/S0266267104001270 Published online by Cambridge University Press

is a natural element of social choice, the desirability of indifference transitivity is perhaps less obvious. Suppose that distribution $(l_1, \ldots, l_n)$ is considered equally good as distribution $(l'_1, \ldots, l'_n)$, due to the importance of an equal value of life principle, and similarly with $(l_1, \ldots, l_n)$ and $(l''_1, \ldots, l''_n)$. Equal value of life reasoning may then not directly suggest that $(l'_1, \ldots, l'_n)$ and $(l''_1, \ldots, l''_n)$ are equally desirable.

Transitivity of indifference may have considerable strength. This is, for example, evident for the situation of Arrowian social choice, for which Sen (1969, 1970) showed that the negative conclusion of Arrow's impossibility theorem is avoided if transitivity of indifference for the social preference relation is relaxed. Although the model here is quite different, we indicate a similar point in the context of making priorities over life profiles.

A binary relation $\succsim$ is *quasi-transitive* if its strict component $\succ$ is transitive. Without indifference transitivity, we do not have an *ordering* of all conceivable distributions. For social choice, however, it may suffice to point to the desired distribution(s) from a given set of feasible distributions. Technically, what we need is a binary relation for which there is a non-empty set of maximal elements for any finite set of alternatives. For this purpose, indifference transitivity is not required.[8] We find that if social choice is interpreted as a selection in accordance with a complete and quasi-transitive binary relation, the equal value of life principle may be consistent with the strong Pareto principle. The proof is by example.

**Proposition 3.** *There exists a complete and quasi-transitive social preference relation satisfying the equal value of life principle and the strong Pareto principle.*

Proof. Consider the following social welfare ordering which by construction satisfies the strong Pareto principle: For any $(l_1, \ldots, l_n)$ and $(l'_1, \ldots, l'_n)$ we have $(l_1, \ldots, l_n) \succ (l'_1, \ldots, l'_n)$ if $l_i \succsim_i l'_i$ for all $i$ and $l_i \succ_i l'_i$ for at least one $i$ and $(l_1, \ldots, l_n) \sim (l'_1, \ldots, l'_n)$ otherwise. Clearly, the social welfare ordering $\succsim$ is complete and since strict individual preference $\succ_i$ is transitive, $\succ$ is also transitive, i.e. $\succsim$ is quasi-transitive. Moreover, the equal value of life principle holds, since for any distribution $(l_1, \ldots, l_n)$ with lifetimes $(t_1, \ldots, t_n)$ and $i \neq j$, if $l_i^{t_i+s} \succ_i l_i$ and $l_j^{t_j+s} \succ_j l_j$, $(l_i^{t_i+s}, l_{k\in N\setminus\{i\}})$ does not weakly dominate and is not weakly dominated by $(l_j^{t_j+s}, l_{k\in N\setminus\{j\}})$ with respect to individual preferences. Thus $(l_i^{t_i+s}, l_{k\in N\setminus\{i\}}) \sim (l_j^{t_j+s}, l_{k\in N\setminus\{j\}})$, which is the equal value of life principle. ■

## 7. CONCLUSION

In this paper we have shown that a priority rule based on the principle of equal value of life or age-dependent equal value of life violates the

[8] Transitivity of strict preference $\succ$ implies acyclicity which is necessary and sufficient for existence of maximal elements on every finite set of alternatives.

weak Pareto principle, if these principles are applied under standard assumptions for social welfare assessment. In other words, health care resource allocation based on principles of equal value of life seems to be in direct conflict with rational social choice. However, for social choice with non-transitive indifference equal value of life principles may be reconciled with this requirement. The Pareto-extension rule used for the proof of Proposition 3 satisfies both (strong) Pareto and (full) equal value of life, but this is obtained by imposing social indifference between any two social states unless one of them weakly dominates the other with respect to the individuals' own preferences. This means that equal value of life is obtained from being only concerned with Pareto optimality and abstaining from any kind of priority setting.

The question is to what extent this normative framework will be acceptable and satisfactory to those defending the equal value of life principle. In any case, our point is that an assessment of this principle depends heavily on what is understood by social preferences and social choice rules. Unambiguity about underlying assumptions seems imperative.

Another question pertains to the general public and policy-makers' positions on these principles. This empirical question remains open for further debate and investigation.

## REFERENCES


Arrhenius, G. 2000. An impossibility theorem for welfarist axiologies. *Economics and Philosophy*, 16:247–66

Cohen, G. 1995. The Pareto argument for inequality. *Social Philosophy and Policy*, 12:160–85

Culyer, A. J. 1997. The rationing debate: Maximising the health of the whole community. The case for. *British Medical Journal*, 314:667–9

Culyer, A. J. and A. Wagstaff. 1993. Equity and equality in health and health care. *Journal of Health Economics*, 12:431–57

Harris, J. 1985. *The Value of Life*. Routledge

Harris, J. 1987. QALYfying the value of life. *Journal of Medical Ethics*, 13:117–23

Harris, J. 1995. Double jeopardy and the veil of ignorance – a reply. *Journal of Medical Ethics*, 21:151–7

Harris, J. 1996. What is the good of health care? *Bioethics*, 10:269–91

Harris, J. 1997. The rationing debate: maximising the health of the whole community. The case against: what the principal objective of the NHS should really be. *British Medical Journal*, 314:669–72

Harris, J. 1999. Justice and Equal Opportunity in Health Care. *Bioethics*, 13:392–404

Johannesson, M. 2001. Should we aggregate relative or absolute changes in QALYs? *Health Economics*, 10:573–7

Kaplow, L. and S. Shavell. 2001. Any non-welfarist method of policy assessment violates the Pareto principle. *Journal of Political Economy*, 109:2:281–86

McKie, J., J. Richardson, P. Singer and H. Kuhse. 1998. *The Allocation of Health Care Resources: An Ethical Evaluation of the 'QALY' Approach*. Ashgate Publishing Co.

Nord, E. 2001. The desirability of a condition versus the well-being and worth of a person. *Health Economics*, 10:579–81


https://doi.org/10.1017/S0266267104001270 Published online by Cambridge University Press

Savulescu, J. 1995. Rational non-interventional paternalism: why doctors ought to make judgements of what is best for their patients. *Journal of Medical Ethics*, 21:327–31

Savulescu, J. 1998. Consequentialism, reasons, value and justice. *Bioethics*, 12:212–35

Savulescu, J. 1999. Desire-based and value-based normative reasons. *Bioethics*, 13:405–13

Sen, A. 1969. Quasi-transitivity, rational choice and collective decisions. *Review of Economic Studies*, 36:381–93

Sen, A. 1970. *Collective Choice and Social Welfare*. Holden-Day

Shaw, P. 1999. The Pareto argument and inequality. *The Philosophical Quarterly*, 49:196:353–68

Singer, P., J. McKie, H. Kuhse and J. Richardson. 1995. Double jeopardy and the use of QALYs in health care allocation. *Journal of Medical Ethics*, 21:144–50

Williams, A. 1997. The rationing debate: Rationing health care by age. The case for. *British Medical Journal*, 314: 820–2

Williams, A. 2001. The 'fair innings argument' deserves a fairer hearing! Comments by Alan Williams on Nord and Johannesson. *Health Economics*, 10:583–5